DOCUMENT RESUME

ED 456 136 TM 033 195

AUTHOR	Pellegrino, James W.
TITLE	Rethinking and Redesigning Education Assessment. Preschool through Postsecondary.
INSTITUTION	Education Commission of the States, Denver, CO.
SPONS AGENCY	Pew Charitable Trusts, Philadelphia, PA.; Metropolitan Life Foundation.
PUB DATE	2001-04-00
NOTE	11p.; Part of a series of essays supported by the Metropolitan Life Foundation Change in Education Initiative.
AVAILABLE FROM	For full text: http://www.ecs.org.
PUB TYPE	Opinion Papers (120)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Computer Assisted Testing; *Educational Technology; Elementary Secondary Education; Models; *Student Evaluation; *Test Construction

ABSTRACT

The recent National Research Council (NRC) report, "Knowing What Students Know: The Science and Design of Educational Assessment," suggests that it is time to rethink the basic assumptions underlying assessment of students and the use of measurement data to enhance teaching and learning. This essay draws on arguments developed in the NRC report to consider why change is needed and how contemporary knowledge and information technologies can bring about the necessary and desired redesign of educational assessment. Expectations about what students should learn have changed in response to social, economic, and technological changes and as a result of the standards-based reform movement. This has resulted in increased testing, although current assessment systems are generally the product of outmoded theories and are characterized by the underuse of technology. Merging research on cognition, measurement, and technology suggests directions for revamping assessment practices to use assessments to inform teachers about the nature of student learning. Powerful information technologies will become as ubiquitous in education as they are in other settings, especially in the design of learning environments. Developing and implementing assessments that take advantage of new technologies will be expensive, but these costs represent an investment in education and the future. (SLD)

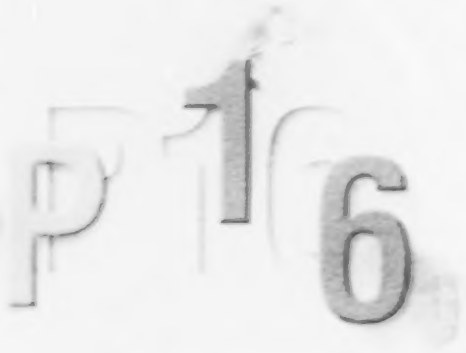
Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 456 136

Rethinking and Redesigning Education Assessment

By James W. Pellegrino

Part of a Series of Essays Supported by the Metropolitan Life Foundation Change in Education Initiative



Preschool Through Postsecondary

TM033195

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

S. F. Walker

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

April 2001

Education Commission of the States • 707 17th Street, Suite 2700 • Denver, CO 80202-3427 • 303-299-3600 • fax 303-296-8332 • www.ecs.org

BEST COPY AVAILABLE

From teachers' informal quizzes to nationally administered standardized tests, assessments have long been an integral part of the educational process. In an ideal system, assessments help teachers, students and parents determine how well students are learning. They help teachers understand how to adapt instruction on the basis of evidence of student learning. They help principals and superintendents document the progress of individual students, classrooms and schools. And they help policymakers and the public gauge the effectiveness of education systems.

With the movement over the past two decades toward setting challenging academic standards and measuring students' progress in meeting those standards, assessment is playing a greater role in decisionmaking than ever before. In turn, education stakeholders are questioning whether current large-scale assessment practices are yielding the most useful kinds of information for informing and improving education. At the same time, classroom assessments, which have the potential to enhance instruction and learning, are not being used to their fullest potential.

The recent National Research Council (NRC) report, *Knowing What Students Know: The Science and Design of Educational Assessment*, suggests that it is time to rethink the basic assumptions underlying how we assess students, and how we use assessment data to enhance teaching and learning. The report points up the critical importance of developing new kinds of classroom and large-scale assessments that work together to help all students learn and succeed in school by making as clear as possible to them, their teachers and other education stakeholders the nature of their accomplishments and the progress of their learning.

This essay draws on arguments developed in the NRC report. Its purpose is to consider why change is needed and how contemporary knowledge and information technologies can bring about the necessary and desired redesign of educational assessment.

Rising Expectations

Expectations about what all students should learn – and, by implication, what they should be tested on – have changed in response to social, economic and technological changes and as a result of the standards-based reform movement. All students are now expected to demonstrate the kinds of reasoning and problem-solving abilities once expected of only a minority of young people. Assessments must therefore tap a broader range of competencies than in the past. They must capture the more complex skills and deeper content knowledge reflected in new expectations for learning. They must accurately measure higher levels of achievement, while also providing meaningful information about students who still perform below expectations. These trends and pressures are being played out on a large scale and accompany the drive to set and meet challenging standards for student learning.

Increased Testing

Standards-based reform has increased both the amount of testing and the stakes attached to test results. Currently, 48 states have statewide testing programs, compared with 39 in 1996, and many school districts also have their own local testing programs (in addition to the range of classroom tests teachers regularly administer). A result of this heightened emphasis on assessment as an instrument of reform is that spending on large-scale testing has doubled in the past four years, from $165 million in 1996 to $330 million in 2000 (Achieve, 2000). More pressure is placed on current assessment systems than they were meant to bear, highlighting several of their limitations.

At least four sets of concerns exist about the current assessment systems of many states:

- **Effectiveness of measurement.** Do the most widely used assessments effectively capture the complex knowledge and skills emphasized in contemporary standards and deemed essential for success in the information-based economy? Probably not. Limits on the kinds of competencies currently being assessed also raise questions about the inferences one can therefore draw from test results. If scores go up on a test that measures a relatively narrow range of knowledge and skills, does that mean student learning has improved, or has instruction simply adapted to a constrained set of outcomes? If there is explicit "teaching to the test," at what cost do such gains in test scores accrue relative to acquiring other aspects of knowledge and skill that are valued in today's society? This is a

point of considerable controversy with regard to the so-called "miracle in Texas" but also for the periodic ups and downs in state assessment results more generally.

- **Utility for improving teaching and learning.** How useful are current assessments for improving teaching and learning – the ultimate goal of education reforms? Not very. Most current large-scale tests provide very limited information that teachers and education administrators can use to identify why students do not perform well, or to modify the conditions of instruction in ways likely to improve student achievement. The most widely used state and district assessments provide only general information about where a student stands relative to peers or whether the student has performed poorly or well in certain domains (for example, that the student performs "below basic" in mathematics). Such tests do not reveal whether students are using misguided strategies to solve problems or fail to understand key concepts within the subject matter being tested. They do not show whether a student is advancing toward competence or is stuck at a partial understanding of a topic that could seriously impede future learning. In short, many current assessments do not provide strong clues as to the types of education interventions that would improve learners' performance, or even provide information on precisely where the students' strengths and weaknesses lie. Nor is information provided in a timely manner.

- **"Snapshots" versus progression over time.** Can we tell how much a student has progressed in a year? Not really. Most assessments provide "snapshots" of achievement at particular points in time, but they do not capture the progression of students' conceptual understanding over time, which is at the heart of learning. This limitation exists largely because most current modes of assessment lack an underlying theoretical framework of how student understanding in a content area develops over the course of instruction, and predominant measurement methods are not designed to capture such growth.

- **Fairness and equity.** Are tests fair and equitable? Perhaps not. Much attention is given to the issue of test bias – whether differences occur in the performance of various groups for reasons that are irrelevant to the competency the test is intended to measure. Standardized-test items are subjected to judgmental and technical reviews to monitor for this kind of bias. The use of assessments for high-stakes decisions, however, raises additional questions about fairness. If the assessments are not aligned with what students are being taught, it is not fair to base promotion or rewards on the results, especially if less advantaged students are harmed disproportionately by the outcome.

If current assessments do not effectively measure the impact of instruction or if they fail to capture important skills and knowledge, how can educators interpret and address gaps in student achievement? One of the main goals of current reforms is to improve learning for low-achieving students. If this goal is to be accomplished, assessment must give students, teachers, administrators and other stakeholders information they can use to improve learning and inform instructional decisions for individuals and groups, especially those not performing at high levels.

Outmoded Theories and Underutilized Technologies

Whether we realize it or not, every educational assessment, whether used in the classroom or large-scale policy context, is based on a set of scientific principles and philosophical assumptions. First, every assessment is grounded in a conception or theory about how people learn, what people know, and how knowledge and understanding progress over time. Second, each assessment embodies certain assumptions about which kinds of observations, or tasks, are most likely to elicit demonstrations of important knowledge and skills from students. Third, every assessment is premised on certain assumptions about how best to interpret the evidence from the observations in order to make meaningful inferences about what students know and can do.

Current assessment systems are the cumulative product of various prior theories of learning and methods of measurement. Although some of these foundations are still useful for certain functions of testing, change is needed. The most common kinds of education tests do a reasonable job with certain limited functions of testing, such as measuring knowledge of basic facts and procedures and producing overall estimates of proficiency for parts of the curriculum. But both their strengths and limitations are a product of their adherence to theories of learning and measurement that are outmoded and fail to capture the breadth and richness of knowledge and competence. The limitations of these theories also compromise the usefulness of the assessments. Assessment systems need to evolve to keep pace with developments in

the sciences of learning and measurement if we are to achieve the learning goals embedded in many of our current standards.

Assessment designs and practices that bring together advances from the cognitive and measurement sciences could address many of the limitations of current assessments and yield a number of benefits for students, teachers and the education system as a whole. Technology is a key factor in bringing about this synthesis. Serious investment in such systems has yet to be made, however.

Rethinking the Foundations of Assessment: The Merger of Cognition, Measurement and Technology

As described in the 1999 NRC reports, *How People Learn: Brain, Mind, Experience, and School*, and *How People Learn: Bridging Research and Practice*, several decades of research in the learning sciences have advanced our knowledge about how children develop understanding in areas of the curriculum, how people reason and build structures of knowledge in academic subject areas, which thinking processes are associated with competent performance, and how knowledge is shaped by social context. For example, studies of expert-novice differences in subject domains have illuminated many critical features of proficiency that should be the targets for assessment. Experts in a subject domain not only "know a lot" – more importantly, they organize knowledge into schemas that support the rapid retrieval and application of such knowledge. Experts also use metacognitive strategies – ways of guiding one's thinking – for monitoring understanding during problem solving and for performing self-correction.

These and many other findings on how people learn and the differences in what novices and experts know suggest directions for revamping assessment practices to move beyond a focus on component skills and discrete bits of knowledge. Assessment should encompass the more complex aspects of student achievement. To aid learning, we need to have access to better information about students' levels of understanding, their thinking strategies and the nature of their misunderstandings.

During the last few decades significant developments also have accrued in measurement methods and theory. A wide array of statistical measurement methods currently are available to support the rigor we want in testing while simultaneously enabling the kinds of inferences about student knowledge that cognitive research suggests are important to pursue when assessing student achievement. In particular, it is now possible to characterize students in terms of multiple aspects of proficiency, rather than a single score; chart students' progress over time, instead of simply measuring performance at a particular point in time; deal with multiple paths or alternative patterns of valued performance; model, monitor and improve judgments based on informed evaluations; and report performance not only at the level of students, but also at the levels of groups, classes, schools and states. Nonetheless, many of the newer models and methods are not widely used because they are not easily understood or packaged in accessible ways for those without a strong technical background.

Technology offers the possibility of addressing this shortcoming. For instance, by building statistical models into technology-based learning environments for use in classrooms, teachers can assign more complex tasks, capture and replay students' performances, share exemplars of competent performance, and in the process gain critical information about student competence. Without question, computer and telecommunications technologies are making it possible to create powerful learning environments and simultaneously assess what students are learning at very fine levels of detail, with vivid simulations of real-world situations, and in ways that are tightly integrated with instruction.

Research has already shown that assessments used to inform teachers about the nature of student learning can help them provide better feedback to students, which in turn significantly can enhance learning. Many of the most effective examples of the use of assessment to inform learning and instruction in the classroom rely on technology-based task presentation and information management systems.

If well-designed and used properly, assessments based on contemporary scientific knowledge also could promote more equitable opportunity to learn by providing better quality information about the impact of education interventions on children. More informative classroom assessments could result in earlier identification of learning problems and intervention for children at risk of failure, rather than waiting for results from large-scale assessments to signal problems. Students with disabilities also could benefit from

this approach. At the same time, it is necessary for educators and researchers to continuously monitor the effects of their practices to ensure that the new assessments do not exacerbate existing inequalities.

Assessments based on contemporary theories and data on how competence develops across grade levels in a curriculum domain also could provide more valid measures of growth and the value added by teachers and schools. Such assessments also could enhance community dialogue about goals for student learning and important indicators of achievement at various grade levels and in different subject areas. Comparisons based on attainment of worthwhile learning goals, rather than normative descriptions of how students perform, could enhance the public's understanding of education quality. New forms of assessment also could help provide descriptive and accurate information about the nature of achievement in a subject area and patterns of students' strengths and weaknesses that would be more useful than existing data for guiding policy decisions and reform efforts.

It is no surprise, then, that collective advances in the study of thinking and learning, in the field of measurement and in the deployment of powerful technologies for learning have stimulated many people to think in new ways about educational futures. New information technologies provide substantial opportunities to advance the design and use of assessments based on a merger of contemporary scientific knowledge of cognition and measurement. Focus is needed on ways to bring together the knowledge of how students learn, what they know and what is therefore worth assessing, with knowledge of how to do this with technical rigor, and ways to harness technology to make the merger feasible. Several intriguing implications arise from projecting what could happen from the coupling of advances in cognition, measurement and technology.

Visions of the Future

Within the next decade, extremely powerful information technologies will become as ubiquitous in education settings as they are in other aspects of people's daily lives. They are almost certain to provoke fundamental changes in learning environments at all levels of the education system. Indeed, reports by groups such as the President's Council of Advisers on Science and Technology and the Web-Based Education Commission indicate that many of these changes are already occurring, prompting conjecture about the consequences for children, teachers, policymakers and the public. Many of the implications of technology are beyond people's speculative capacity. A decade ago, for example, few could have predicted the sweeping effects of the Internet on education and other segments of society. The range of computational devices and their applications is expanding exponentially, fundamentally changing how people think about communication, connectivity, information systems, education practices and technology's role in society.

While it is always risky to predict the future, it appears clear that advances in technology will continue to impact the world of education in powerful and provocative ways. Many technology-driven advances in the design of learning environments, which include the integration of assessment with instruction, will continue to emerge and will reshape the terrain of what is both possible and desirable in education. Advances in curriculum, instruction, assessment and technology are likely to continue to move education practice toward a more individualized and mastery-oriented approach to learning. This evolution will occur across the K-16 spectrum. To manage learning and instruction effectively, people will want and need to know considerably more about what has been mastered, at what level and by whom.

Consider the possibilities that might arise if assessment is integrated into instruction in multiple curricular areas and the resultant information about student accomplishment and understanding is collected with the aid of technology. In such a world, programs of on-demand external assessment such as state achievement tests might not be necessary. Instead, it might be possible to extract the information needed for summative and program evaluation purposes from data about student performance continuously available both in and out of the school context.

Technology could offer ways of creating, over time, a complex stream of data about how students think and reason while engaged in important learning activities. Information for assessment purposes could be extracted from this stream and used to serve both classroom and external assessment needs, including providing individual feedback to students for reflection about their learning strategies and habits. To realize this vision, additional research on the data representations and analysis methods best suited for

different audiences and different assessment objectives clearly would be needed – and is certainly doable.

We can therefore imagine a future in which the audit function of assessments external to the classroom would be reduced significantly or even unnecessary because the information needed to assess students, at the levels of description appropriate for various monitoring purposes, could be derived from the data streams generated by students in and out of their classrooms.

A metaphor for such a radical shift in how one "does the business of educational assessment" exists in the world of retail outlets, ranging from small businesses to supermarkets to department stores. No longer do these businesses have to close down once or twice a year to take inventory of their stock. Rather, with the advent of automated checkouts and barcodes for all items, these enterprises have access to a continuous stream of information that can be used to monitor inventory and the flow of items. Not only can business continue without interruption, but the information obtained is far richer, enabling stores to monitor trends and aggregate the data into various kinds of summaries. Similarly, with new assessment technologies, schools would no longer have to interrupt the normal instructional process at various times during the year to administer external tests to students. Nor would they have to spend significant amounts of time preparing for specific external tests peripheral to the ongoing activities of teaching and learning.

Extensive technology-based systems that link curriculum, instruction and assessment at the classroom level might enable a shift from today's assessment systems, which use different kinds of assessments for different purposes, to a balanced design in which the three critical features of *comprehensiveness*, *coherence* and *continuity* would be ensured. In such a design, assessments would provide a variety of evidence to support education decisionmaking (*comprehensiveness*). The information provided at differing levels of responsibility and action would be linked back to the same underlying conceptual model of student learning (*coherence*) and would provide indications of student growth over time (*continuity*).

Clearly, technological advances will allow for the attainment of many of the goals that educators, researchers, policymakers, teachers and parents have envisioned for assessment as a viable source of information for educational improvement. When powerful technology-based systems are implemented in classrooms, rich sources of information about student learning continuously will be available across wide segments of the curriculum and for individual learners over extended periods of time. This is exactly the kind of information we now lack, making it difficult to use assessment to truly support learning. The major issue is not whether this type of data collection and information analysis is feasible in the future. Rather, the issue is how the world of education anticipates and embraces this possibility, and how it will explore the resulting options for effectively using assessment information to meet the multiple purposes served by current assessments and, most important, to enhance student learning.

It has been noted that the best way to predict the future is to invent it. Without doubt, multiple futures for educational assessment could be invented on the basis of synergies we know exist among information technologies and contemporary knowledge of cognition and measurement. While we are a considerable distance away from implementing the types of fully integrated systems envisioned above, there are steps to take now that would put us on the path to such a future.

> State policymakers are urged to recognize the limitations of current assessments and to support the development of new systems of multiple assessments that would improve their ability to make decisions about education programs and the allocation of resources.
>
> - Important decisions about individuals should not be based on a single test score. Policymakers should invest instead in the development of assessment systems that use multiple measures of student performance, particularly when high stakes are attached to the results.
> - Assessments at the classroom and large-scale levels should grow out of a shared knowledge base about the nature of learning. Policymakers should support efforts to achieve such coherence.
> - Policymakers should promote the development of assessment systems that measure growth or progress of students and the education system over time, and that support multilevel analyses of the influences responsible for such change.

The current educational assessment environment in the United States reflects the considerable value and credibility placed on external, large-scale assessments of individuals and programs versus classroom assessment designed to assist learning. The resources invested in producing and using large-scale testing in terms of money, instructional time, research and development far outweigh the investment in the design and use of effective classroom assessment. Not only does large-scale assessment dominate over classroom assessment, but there also is ample evidence of accountability measures negatively impacting classroom instruction. Teachers feel pressure to teach to the test, which results in a narrowing of instruction. They also model their own classroom tests after less-than-ideal standardized tests.

These kinds of problems suggest that beyond striking a better balance between classroom and large-scale assessment, what is needed are coordinated systems of assessments that collectively support a common set of learning goals, rather than working at cross-purposes. A range of assessment approaches should be used to provide a variety of evidence to support education decisionmaking.

Ideally in a balanced assessment environment, a single assessment does not function in isolation, but rather within a "nested" assessment system involving states, local school districts, schools and classrooms. Assessment systems should be designed to optimize the credibility and utility of the resulting information for both education decisionmaking and general monitoring. To this end, an assessment system should, as mentioned above, exhibit three critical properties: comprehensiveness, coherence and continuity. These three characteristics describe an assessment system that is aligned along three dimensions: vertically, across levels of the education system; horizontally, across assessment, curriculum and instruction; and temporally, across the course of a student's studies.

Multiple assessments are needed to provide the various types of information required at different levels of the education system. Single measures, while useful, are unlikely to tap all the dimensions of competence identified by learning goals. If multiple assessments grow out of a shared knowledge base about cognition and learning in an academic domain, they can provide valuable multiple perspectives on student achievement while supporting a core set of learning goals.

Stakeholders should not be unduly concerned if differing assessments yield different information about student achievement; in fact, in many circumstances this is exactly what should be expected. If multiple assessments, however, are to support learning effectively and provide clear and meaningful results for various audiences, it is important that the purposes served by each assessment and the aspects of achievement that any given assessment samples be made explicit to multiple audiences.

Assessments within such a system should be aimed at improving learning by providing information needed by those at all levels of the education system on the aspects of schooling for which they are responsible. If properly conducted, assessments also can serve accountability purposes by providing valuable information to teachers and administrators about the progress or growth of the education system over time. And if the assessments are instructionally sensitive – that is, if they show the effects of high-quality teaching – they can provide important information about the effectiveness of teaching practices as well.

Developing and implementing a system of multiple assessments would likely be more costly than continuing with the array of tests now being used by states and school districts. While the $330 million a year that states currently spend for testing appears considerable, it represents less than one-tenth of 1% of the total amount spent on precollege education (National Center for Education Statistics, 2001). If used properly, the total spending for assessment should not be considered money for tests alone. Funds spent for teachers to score assessments, included in the cost of assessment, also serve an important professional development function. Moreover, spending on assessments that inform instruction represents an investment in teaching and learning, not just in system monitoring. Policymakers, therefore, need to invest considerably more in assessment than is currently the case, presuming that the investment is in assessment systems of the type advocated in the NRC report *Knowing What Students Know: The Science and Design of Educational Assessment*.

> Leverage should be exerted on developers of educational curricula and materials to create tools that will enable teachers to implement high-quality instructional and assessment practices, consistent with modern understanding of how students learn and how such learning

can be measured.

- Assessments and supporting instructional materials should be developed that interpret the findings from cognitive research in ways that are useful for teachers.
- Advantage should be taken of opportunities that technology affords to assess what students are learning, at fine levels of detail, with appropriate frequency and in ways that are tightly integrated with instruction.

Assessment based on a synthesis of contemporary cognitive and measurement principles particularly has significant potential for the design of high-quality tools for classroom assessment that can inform and improve learning. Several innovative classroom assessments have emerged that are having a positive impact on learning, but much more needs to be done to promote their development and use. A key to the effectiveness of these tools is that they must be packaged in practical ways for use by teachers. In this regard, computer and telecommunications technologies offer a rich array of opportunities for providing teachers with sophisticated assessment tools that allow them to gain critical information about student competence.

Assessment tasks no longer need be confined to paper-and-pencil formats, and the entire burden of classroom assessment need no longer fall on the teacher. Among the most intriguing applications of technology are those that extend the nature of the problems that can be presented and the knowledge and cognitive processes that can be assessed. By enriching task environments through the use of multimedia, interactivity and control over the stimulus display, it is possible to assess a much wider array of cognitive competencies than has been feasible heretofore. New capabilities enabled by technology include directly assessing problem-solving skills, making visible sequences of actions taken by learners in solving problems, and modeling and simulating complex reasoning tasks. Technology also makes possible data collection on concept organization and other aspects of students' knowledge structures, as well as representations of their participation in discussions and group projects.

A significant contribution of technology has been to the design of systems for implementing sophisticated classroom-based formative assessment practices. Technology-based systems have been developed to support individualized instruction by extracting key features of learners' responses, analyzing patterns of correct and incorrect reasoning, and providing rapid and informative feedback to both student and teacher.

Technology also is having an influence on what is taught and how. Examples include the teaching of advanced thinking and reasoning skills within a discipline through the use of technology-mediated projects involving long-term inquiry. Such projects often integrate content and learning across disciplines, and show how technology makes possible the integration of assessment with curriculum and instruction in powerful ways. As noted earlier, increased availability of such systems could make it possible to pursue balanced designs representing a more coordinated and coherent assessment system. Information from such assessments could possibly be used for multiple purposes, including the audit and accountability function associated with many existing large-scale assessments.

Resources should be invested in the development of large-scale assessments that sample the broad range of competencies and forms of student understanding that research shows are important aspects of student learning.

- A variety of matrix sampling, curriculum-embedded and other assessment approaches should be used to cover the breadth of cognitive competencies that are the goals of learning in a domain of the curriculum.
- Large-scale assessment tools and supporting instructional materials should be developed so that clear learning goals and landmark performances along the way to competence are shared with teachers, students and other education stakeholders. The knowledge and skills to be assessed and the criteria for judging the desired outcomes should be clearly specified and available to all potential examinees and other concerned individuals.
- New ways of reporting assessment results that convey important differences in performance at various levels of competence – in ways that are clear to different users, including

educators, parents and students – should be pursued.

Though further removed from day-to-day instruction than classroom assessments, large-scale assessments have the potential to support instruction and learning if well designed and appropriately used. Deriving real benefits from the merger of cognitive and measurement theory in large-scale assessment requires finding ways to cover a broad range of competencies and capture rich information about the nature of student understanding. Alternatives to the typical on-demand testing scenario – in which every student takes the same test at a specified time under strictly standardized conditions – should be considered to enable the collection of more diverse evidence of student achievement.

Large-scale assessments have an important role to play in providing dependable information for use by policymakers, school administrators, teachers and parents. Large-scale assessments also can convey powerful messages about the kinds of learning valued by society and provide worthy goals to work toward. If such assessments are to serve these purposes, however, it is essential that externally set goals for learning be clearly communicated with teachers, students and other education stakeholders.

Considerable resources should be devoted to producing materials for teachers and students that clearly present both the learning goals and landmark performances along the way to competence. Those performances can be illustrated with samples of the work of learners at different levels of competence accompanied by explanations of the aspects of cognitive competence exemplified by the work. These kinds of materials can foster valuable dialogue among teachers and students, and the public about what achievement in a specific domain of the curriculum looks like. The criteria by which student work will be judged on an assessment also should be made as explicit as possible. Curriculum materials should encourage the use of activities such as peer- and self-assessment to help students internalize the criteria for high-quality work and foster metacognitive skills.

The use of assessments based on advances in the cognitive and measurement sciences also will require different forms of reporting on student progress, both to parents and to administrators. The information gleaned from such assessments is far more nuanced than that obtainable from the assessments commonly used today, and teachers may want to provide more detail in reports to parents about the nature of their children's understanding.

In formulating reports based on new assessments, test developers, teachers and school administrators should ensure that the reports include the information parents want and can appropriately use to support their children's learning. Reports on student performance also could provide an important tool to assist administrators in their supervisory roles. Administrators could use such information to see how teachers are gauging their students' learning and how they are responding to students' demonstration of understanding. Such information could help administrators determine where to focus resources for professional development. In general, for the information to be useful and meaningful, it will have to include a profile consisting of multiple elements and not just a single aggregate score.

Concluding Comment

Although periodic testing and assessment is a critical part of any education reform and accountability process, much of the movement toward increased testing for accountability has been fueled by a misguided assumption that more frequent testing, in and of itself, will improve education. There is a "down home" expression that captures the essence of the problem: "You don't fatten a pig by constantly weighing it. Eventually, you have to feed it if you want it to grow."

At the same time, criticism of education accountability policies may be predicated on an equally misguided assumption that testing, in and of itself, is responsible for most of the problems in education. This is wishful thinking on the part of many. A more realistic view is to address our ongoing concerns about education quality and improvement not by stepping up the amount of testing or abandoning assessments entirely, but rather by refashioning assessments in principled ways to meet current and future needs for quality information about what students know.

James W. Pellegrino is the Frank W. Mayborn Professor of Cognitive Studies at Vanderbilt University. Previously, he served as dean of Vanderbilt's Peabody College of Education and Human Development

and as co-director of Vanderbilt's Learning Technology Center. He has authored or co-authored numerous books, chapters and journal articles in the areas of cognition, instruction and assessment.

This is one in a series of papers aimed at stimulating dialogue and action on the need for state-level system redesign in American public education. With the support of Metropolitan Life Foundation and the Pew Charitable Trusts, nine reports will be published during 2001. The reports will be available on the ECS Web site at www.ecs.org.

© Copyright 2001 by the Education Commission of the States (ECS). All rights reserved.

The Education Commission of the States is a nonprofit, nationwide organization that helps state leaders shape education policy. It is ECS policy to take affirmative action to prevent discrimination in its policies, programs and employment practices.

To request permission to excerpt part of this publication, either in print or electronically, please fax a request to the attention of the ECS Communications Department, 303-296-8332 or e-mail ecs@ecs.org.

BEST COPY AVAILABLE